

AD-A070 820

NAVAL HEALTH RESEARCH CENTER SAN DIEGO CA
A CASE FOR PSYCHIATRIC DIAGNOSIS.(U)
DEC 74 J S FICHMAN, D EDWARDS
NAVHLTHRSCHC-75-10

F/G 5/10

UNCLASSIFIED

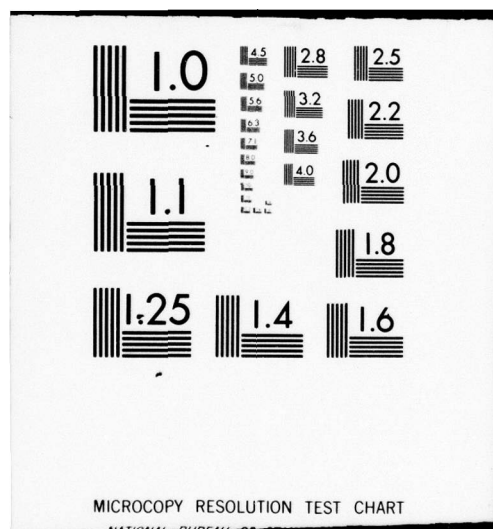
NL

| OF |
AD
A070 820



END
DATE
FILMED
8-79

DDC



LEVEL 11

2
B.S.

A CASE FOR PSYCHIATRIC DIAGNOSIS

A070820

10
J. S. FICHMAN
D. EDWARDS

REPORT NO. 75-10

DDC
REFILED
JUL 3 1979
RECEIVED

DDC FILE COPY



This document has been approved
for public release and sale; its
distribution is unlimited.

NAVAL HEALTH RESEARCH CENTER

P.O. Box 86122
San Diego, CA 92161-0152

NAVAL MEDICAL RESEARCH AND DEVELOPMENT COMMAND

BETHESDA, MARYLAND

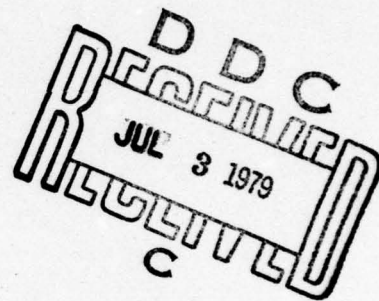
B

A CASE FOR PSYCHIATRIC DIAGNOSIS ¹

J. Susan Fichman

and

Darrel Edwards



Naval Health Research Center

San Diego, California

¹Report No. 75-10, supported under research work unit M4305.07-3013DXB5.

Opinions expressed are those of the author and are not to be construed as necessarily reflecting the views or endorsement of the Navy Department.

ABSTRACT

Psychiatric diagnosis has been severely criticized by many clinicians and researchers. Studies in clinical judgment and diagnosis leading to this criticism were reviewed. Research designs were evaluated in terms of their relevance to diagnosis. Most projects were extremely narrow in scope and it was inappropriate to apply their conclusions to the process of diagnosis. Research evaluating diagnosis in real clinical settings showed that diagnosis accurately and consistently assessed the patients' capacity for successful post-hospitalization readjustment. In an effort to evaluate diagnosis as practiced in the clinical setting, the grid method was used. Psychiatrists rated patients from their practice, generating their own dimensions for comparison. Patients within a diagnostic category were consistently evaluated as being more similar to each other than patients from different diagnostic categories. The use of this highly individual technique assured that even analysis of grouped data would not obscure the basic strength of the diagnostic process.

Accession For	
NRIC - GRII	<input checked="checked" type="checkbox"/>
DOC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By	
Distribution/	
Availability Codes	
Dist	Available and/or special
A	

A CASE FOR PSYCHIATRIC DIAGNOSIS

J. Susan Fichman and Darrel Edwards

Naval Health Research Center
San Diego, California

Psychiatric diagnosis has been under strong attack for the last ten years. Research over the last 20 years has produced both positive and negative conclusions about the validity and reliability of judgment and diagnosis. This wealth of obscurity has lead many clinicians and researchers to propose abandoning the entire psychiatric diagnostic system (Kanfer & Saslow, 1969; Frank, 1969). Against the move for abandonment is a body of research which established the utility of diagnosis in clinical settings (Berry & Edwards, 1974; Edwards, Fichman, Bucky & Berry, 1974;) Fichman, Edwards & Berry, 1973). These studies have clearly shown that psychiatric diagnosis is the best predictor of occupational performance after treatment

This study reviews the literature in an effort to understand the difference between the positive and negative results reported in the past. In addition, a proposal is made for a method to measure the diagnostic process in actual clinical practice before the decision to abandon the present diagnostic system is made.

Two general approaches have been used in evaluating clinical judgment and diagnosis. First, experimental paradigms have sought to answer the following questions: (a) Does psychiatric experience relate to accuracy of judgment? (b) Do differing amounts and types of information relate to accuracy? (c) Do

Judges agree with each other and with their own previous judgments? Secondly, descriptive and model-generating studies have been initiated in an effort to understand the clinician's judgmental processes. In this study, research on these two approaches will be reviewed and a third approach will be developed.

Professional Experience and Accuracy

Increased education and experience in psychiatry and psychology were assumed to be related to increased accuracy in diagnostic evaluation. In examining the relationship between accuracy and experience, accuracy was assessed by the clinician's ability to predict patient response to personality questionnaires, Q-Sorts, adjective check lists, or vocabulary test items from varying types of minimal information. In determining the effect of experience on clinical judgment, few studies have tried to replicate diagnosis in clinical settings. Instead, experimenters have parceled out bits of information and asked clinicians to predict responses to various types of tests. Experiments conducted within this restricted paradigm have yielded both positive and negative results.

In 1950 Luft asked a group of clinicians and a group of physical scientists to listen to a case conference and predict that patient's responses to a personality questionnaire. The hypothesis that the clinicians with their specialized experience and training would be more accurate than their equally educated but psychologically untrained counterparts, was not supported. The patient's therapist also filled out the personality questionnaire; his responses were more accurate than ninety-seven percent of the judges. This finding was not statistically or qualitatively evaluated by Luft. But it seems to point up the gap between the elements of the experimental design and the content of the

real clinical situation (Luft, 1950)

In an effort to examine the practical mechanism of clinical judgment, Gunderson (1965, a & b) surveyed clinicians involved in Operation Deep Freeze. Pairs of clinicians rated men for special duty in the Antarctic. Complex patterns generated to describe the judgmental process explained relatively little variance. It was hypothesized that to account for more of the variance, it would be necessary to know the meaning for the clinician of the various traits. This hypothesis indicated a need to uncover the personal constructs used in evaluation.

Oskamp (1962) examined the effect of clinical experience on predictive ability. Clinicians and undergraduate psychology majors evaluated 200 MMPI profiles. One hundred profiles were from psychiatric patients and 100 were from people with no psychiatric disturbance. These profiles were scored to indicate in which group they belonged and the judge estimated his confidence in his decision. Experience was positively related to accuracy, but negatively related to confidence.

In an experiment where judges had more varied levels of experience, Grigg (1958) found no difference between graduate student trainees and clinical psychologists in predicting responses to Gough's adjective check list and a self-report questionnaire from interview data. However, those judges with one or more years experience were significantly more accurate than naive judges.

Hunt, Jones and Hunt (1957) asked trained clinical psychologists and beginning psychology students to rate schizophrenic responses to vocabulary test items. They found no difference between the means of the two groups, but there was a significant difference between the standard deviations. The trained clinicians showed less dispersion in their ratings indicating more inter-rater reliability. Johnston and McNeal (1967) while examining statistical and

clinical prediction, found that neither professional background nor years of experience was related to accuracy.

Grigg's study was typical of the work evaluating the effect of experience on clinical judgment. Asking clinicians to predict responses to personality questionnaires from various bits of data was a very tenuous representation of the diagnostic procedure, both in information available and task required. The exceptionally high accuracy of the patient's therapist reported by Luft (1950) pointed up the lack of congruency between experimental manipulations and clinical functioning. Luft ignored the individual therapist and was only concerned with the results from his grouped data. Gunderson (1965, a & b) also used a nomothetic analysis but he advised that an understanding of the individual clinician's system of evaluation would provide a more productive approach to the examination of clinical judgment. To gain insight into the process of diagnosis, it is essential to understand the individual and use unobtrusive measures taken in the occupational setting. The more important evaluation of clinical experience is whether it enables clinicians to make accurate diagnoses and treatment decisions and whether these decisions relate positively to success in post-treatment adjustment.

In examining the effect of clinical experience, investigators have also asked if attributes of the clinician systematically influence their judgment. Does clinical judgment reflect patient illness severity or intra-clinician variables? Harrison, McDermott, Schager and Showerman (1970) examined the influence of the psychiatrist's class background on diagnosis, clinical perception, prognosis, and recommendation for treatment. Their results showed that psychiatrists with lower class backgrounds gave diagnoses of psychosis and personality disorder more frequently and neurosis less frequently than upper class evaluators.

When investigating the interaction of social class and attitudes of the patient and interviewer, as they related to diagnosis and disposition, Shader and Binstock (1969) found that the social class of the patient and the attitudes of the diagnostician were significantly related to diagnosis and disposition. Patients of a lower social class were more likely to be diagnosed psychotic. If the diagnostician liked the patient he was more likely to be diagnosed neurotic and to receive therapy at the outpatient facility. In investigating this same area, Lowinger and Dobie (1968) discovered that psychiatric attitudes were related to diagnosis, race, religion, social class, sex, marital status, and age of the patient. The number of outpatient appointments was influenced by the patient's age, sex, social class, and miscellaneous attitudinal factors. Choice of psychotherapy or drug treatment was related to patient social class and attitude factors.

Several authors have presented comprehensive reviews of the literature on personal variables associated with the clinician and his client (Garfield, 1971; Meltzoff and Kornreich, 1970). The only clear finding was that many personal characteristics of the clinician and his client correlated with all phases of clinical judgment. This confusing picture pointed up the need for idiographic measures to understand clinical functioning. The idiographic method of focusing on the individual using techniques and variables that are relevant to the uniqueness of that person is pertinent to this problem. Katz, Cole, and Lowery (1969) concluded that disagreement among clinicians might have been due to actual differences in their perceptions of certain kinds of pathology rather than in semantic preference. This conclusion focused on the importance of examining individual behavior if clinical judgment was to be understood. It is important to discern what these perceptual differences are and how they influence diagnostic accuracy. Knowledge of personal functioning should be the

basis for the evaluation of parameters relevant to more efficient performance.

Differential Information and Accuracy

In examining clinical judgment several experimenters have manipulated the amount and type of information available to the clinician to ascertain its effect on clinical accuracy and reliability. The hypothesis tested was that increasing amounts and types of information would result in increasing accuracy. The experimental condition usually established was that judges were asked to predict patient test responses from varying types and amounts of information. The background information was typically from projective tests or case histories. In background information supplied and projective evaluations requested this design was similar to those used for evaluating the effect of clinical experience.

Kostlan (1954) asked experienced psychologists to evaluate with a true or false response, statements about a client on the basis of different amounts of information. There were five possible pieces of information: Rorschach, MMPI, sentence completion test, social case history, and minimal data.(name and age only). Psychologists evaluated clients with all information except Rorschach, or MMPI, or sentence completion, or they were only given the minimal data. Minimal data produced results no better than chance. Information without social history produced results no better than chance. Superior batteries contained the MMPI and social case history.

Golden (1967) established two groups of judges--criterion judges and test judges. Criterion judges reviewed extensive patient case histories and then completed a personality questionnaire on five heterogeneous subjects. Test judges were asked to complete the personality questionnaire on the basis of identifying data alone, MMPI alone, TAT alone, Rorschach alone or in pairs involving all possible permutations, or all four combined. Concurrent validity was defined as agreement between criterion judges and test judges. Reliability

was defined as agreement between test judges. Reliability and validity did not increase as a function of the number of tests nor were there any differences between tests and pairs of tests. Hunt and Walker (1966) combined protocols from the vocabulary and comprehension sections of the WAIS. The combination did not result in a better diagnosis than when clinicians used either section alone. In another experiment using combinations of the WAIS alone, the WAIS and Beta profile, or the WAIS and figure drawing, Huff and Freidman (1967) found that redundant information, WAIS with Beta profile, reduced reliability. But new information, WAIS and figure drawing versus WAIS alone increased reliability. Oskamp (1965) discovered that increasing information drawn from sections of a published case study resulted in no significant increase in accuracy of personality judgment. But confidence increased significantly and steadily with increases in information.

Goldberg and Werts (1966) evaluated various tests and a vocational history. Experienced clinicians ranked each of four samples of ten patients on one of four traits using one of four sources of information. The traits ranked were social adjustment, ego strength, intelligence, and dependency. The four sources of information were MMPI, Rorschach, Wechsler, and a vocational history. Findings indicated that judgments by one clinician from one data source had no systematic relationship to those of another clinician working from another data source even though they were ranking the same patient. These findings were not surprising in light of extensive review of test validity as it was affected by the requested experimental task (Rapaport, Gill, and Schafer, 1968). Gunderson (1965) examined psychiatrists, who evaluated 719 men on various personality traits. The basis for their evaluation was either Rorschach responses, or interview and biographic data. The results indicated that using different sources of information affected reliability.

Studies evaluating the effect of different amounts and kinds of information available on the accuracy of clinical judgments typically forced a prediction task that is uncharacteristic of clinical decisions in occupational settings. Predicting responses to personality questionnaires is a skill different from the ability to correctly assess the patient's level of emotional impairment or ability to function effectively in a working situation. A more viable evaluation of the effect of information should employ a more appropriate criterion. It would also be useful to consider individual differences in effective information usage.

Inter- and Intra-judge Reliability

Another point in examining the stability and usefulness of diagnosis and clinical judgment is the importance of evaluating inter- and intra-judge reliability. To facilitate communication it is crucial that judges, viewing the same patient and having access to the same information, reach similar diagnostic decisions. It is also critical for the individual judge to be operating in a consistent manner rather than functioning randomly. Inter-judge reliability has been more thoroughly scrutinized by experimenters than intra-judge reliability. Information on inter-judge reliability was previously reported in studies dealing with rating stability across data sources, the effect of clinical experience, and evaluation of personal attributes of the clinician (Gunderson, 1965, a & b; Grigg, 1958; Golden, 1967; Hunt, Jones, and Hunt, 1957; Harrison, McDermott, Schrager, and Showerman, 1960).

Foulds (1955) examined psychiatric agreement between institutionally established final psychiatric diagnoses and diagnostic predictions made from a short battery of tests. In one condition the psychiatrist administered and intergrated the tests. In this condition agreement was 4.5 on a scale of 6. In the second condition the psychiatrist made the diagnosis from the tests without seeing the patient; agreement was 3.83. The difference was not

significant. Diagnosis based solely on knowledge of baseline diagnostic categories admitted showed a mean agreement with the criterion of less than 2.

Schmidt and Fonda (1956) correlated diagnoses made by psychiatrists and residents on 426 patients. With general classifications of organic, psychoneurosis, and character and behavior disorder agreement between psychiatrists and residents was $r = .90$. Agreement as to a specific subtype was only accurate in one-half of the cases and agreement was almost absent in psychoneurosis and character and behavior disorders. Specific organic problems appeared to be more easily diagnosed but differential rater experience confounded the interpretation of rater disagreement.

In an experiment designed to evaluate prognostic judgment, twenty-two psychiatrists rated two specific diagnostic categories within functional psychosis as to which diagnosis had the poorer prognosis. Statistically, significant intra-judge and inter-judge agreement was found $p < .01$ (Stone, 1966). Hunt, et al (1957) examined clinical psychologists' and beginning psychology students' ratings of schizophrenic responses to vocabulary test items and found that judges with more training agreed with each other significantly more often than did naive judges.

Phelan (1964) asked twenty psychologists to match sixteen projective and objective documents with individual biographies. The information was presented in arrays with unequal matching. Ten judges matched information to biographies at the .05 level of significance; the other ten judges performed at chance level. It was hypothesized that bad judges differed from good judges in that poor judges projected their personality and needs thus distorting information about the patient. Phelan (1965) again asked judges to match projective and objective test information with autobiographies for six subjects. Judges differed among themselves, but individuals were consistent. A judge, who was superior at

matching a test autobiography, was superior at matching all tests. These two studies emphasize the need for understanding individual styles of clinical judgment.

Phelan (1964, 1965) showed that judges differed markedly in their ability to perform the matching task he requested of them. This difference in individual ability was reported in a clinical setting in the Navy by Edwards, Gunderson, Brown, and Taylor (1973). They studied psychiatrists with high success rates and psychiatrists with low success rates. Success was defined with on-the-job measures. A patient was judged to have been successful if after two years the man had not been re-hospitalized while in the service or if he had been discharged from the service with a recommendation for reenlistment. There were significant differences in the treatment strategies used by the two groups. Phelan's (1964, 1965) experimental findings and Edwards' et al (1973) occupational data suggested that consideration of individual differences and knowledge of personal cognitive styles was important. Data on the accuracy of clinical prediction could be correlated with information on cognitive style to determine which evaluative systems are most efficient.

Descriptive Studies

In order to develop a working understanding of diagnostic labels and the diagnostic process many investigators have done descriptive studies of diagnosis. Gaaron and Dickenson (1966) organized case history information into small units of data. Clinicians were given a list of these units and instructed to request information as they needed it to arrive at a diagnosis. The order and frequency of the requests were recorded. They concluded that ten pieces of information were necessary and sufficient to establish a diagnosis. The necessary pieces of information were: (1) reason for referral, (2) age, (3) previous personality, (4) mental status--content of thought, (5) sex, (6) previous episodes of

illness, (7) projective testing, (8) general appearance, (9) mental status--affect, and (10) mental status--stream of thought.

Mahrer, Thorp, and Sternlicht (1960) randomly selected 120 sentences from social histories, and psychological and psychiatric reports. One sentence units were used as cues. Clinicians were asked to respond to each cue with the diagnosis it brought to mind. They were then asked to weight the degree of certainty attached to that diagnosis. Results showed that the cues were systematically related to sets of diagnostic categories; the correlations were both positive and negative.

Nathan and his associates have examined symptom patterns of various diagnostic categories to ascertain if there were unique configurations of symptoms that distinguished one diagnostic category from another. Each experiment revealed some unique factors for each diagnostic category and several areas of overlapping symptomatology. Depression and anxiety were shown to be universal (Nathan, 1969; Nathan, Robertson, and Andberg, 1969; Nathan, Gould, Zare, and Roth, 1969).

Overlapping symptoms were noted by Zigler and Phillips (1961) when they extracted symptoms from 793 case studies to see if specific symptoms related to specific diagnoses. They found that some patterns emerged, but there was considerable overlap. Symptoms significantly related to one diagnosis were also related to other diagnoses. Their conclusion was that a diagnosis does not represent a clear symptom pattern.

Ellis and Sells (1964) investigated patient's symptomatic information related to official diagnostic nomenclature in a military setting. They found no reliable patterns. Arthur and Gunderson (1966) examined the relationship of demographic data and narrative summaries to diagnosis or disposition. Neither diagnosis nor disposition was highly predictable.

Frank (1969) analyzed the Rorschach, Wechsler, MMPI, and an analysis of behavior to ascertain how they related to diagnosis. There were no significant differences between diagnostic groups for any of these measures. It was suggested that an entirely new system was necessary.

Descriptive studies do the least damage to the elements of clinical functioning, but researchers focused on nomothetic evaluation. It would be useful to investigate diagnosis with a methodology which allows the individual to reveal the cognitive structure he uses to successfully differentiate among people. Knowledge of this structure could then be used to examine patient variables related to various diagnostic labels.

Model Generation

For some investigators, the efficient description of clinical judgment has led to attempts to build models of the process. In several studies, Stone (1968, 1969, 1970) has sought to use a psychophysical approach to develop subjective scales of prognostic favorability as it relates to psychiatric nomenclature and patient variables. He found that prognostic favorability magnitude estimations were related to percentage improvement, length of stay in hospital, and age. These relationships were in the form of a power curve. In 1960, Hoffman attempted to develop mathematical equations to represent judgmental processes. He asked several judges to predict sociability from variables of the Edward's Preference Scale. He then selected one judge to study intensity in order to develop a configural mathematical equation for that judge. The configural formula developed did not predict better than a linear formula.

Beenan, Van Frankenhuisen, and Veldkamp (1972) created a computer simulation of the diagnostic system. The simulation searched for groups of patients who resembled each other most. Categories produced by this method were judged to be clinically meaningful by a psychiatrist. Johnston and McNeal (1967)

compared the accuracy of a statistical and clinical prediction methods in predicting the length of a patient's stay in the hospital. Clinicians achieved a 71.94% accuracy while statistical methods were 71.92% accurate.

In a review of the literature on statistical models of clinical judgment, Goldberg (1968) evaluated linear, configural, and curvilinear models of clinical judgment to determine which model most efficiently and accurately reproduced clinical judgment. Several of the models reviewed were built to represent individual functioning. In most cases, it was found that linear equations reproduced clinical judgments most accurately. For the few clinicians whose functioning was best represented by a configural model the improvements were of very small magnitude. The most unusual feature of much of the work done with models has been the use of an idiographic approach. Discovering how individuals function is a first critical step in understanding the diagnostic process.

Evaluation of Diagnosis in an Occupational Setting

The tightly designed experiments created to answer specific questions lacked the scope necessary to evaluate the ongoing process of diagnosis in the clinical setting. The manipulations employed were germane to the academic questions asked but were irrelevant to an evaluation of the entire diagnostic process. A more relevant test of diagnosis is whether diagnosis communicates some useful information about the patient's ability to function.

In studying the clinician as he functions in naval inpatient and out-patient settings, several studies have found that diagnosis is a powerful predictor of effectiveness and disposition (Edwards, et al, 1973; Fichman, et al, 1973). The findings of Edwards, et al (1973) demonstrated that diagnosis was a powerful predictor of effectiveness ($r = .41$). Effectiveness was defined as completion of current enlistment and recommendation for reenlistment. In a regression analysis with effectiveness as a criterion, years of service ($r = .41$), diagnosis ($r = .41$) and number of days in the hospital ($r = .26$) significantly

contributed to a multiple regression equation of $R = .56$, which cross-validated at $r = .55$. For the analysis diagnosis was divided into the four major categories: psychosis, neurosis, personality disorder and situational maladjustment. The argument was proposed that diagnosis follows disposition. Disposition was defined as the recommendation to return a man to duty or not. To answer this question an analysis was made of those men who were returned to duty after hospitalization. Nine percent of all psychotics, twenty-four percent of personality disorders, thirty-two percent of neurotics, and seventy-nine percent of those diagnosed situational maladjustment were returned to duty. A second regression analysis was run with effectiveness as the criterion. Diagnosis dropped to eighth most powerful predictor ($r = .21$) of eleven significant variables. The multiple R was $.45$ and cross-validated at $r = .40$. Of the patients returned to duty, seventy-nine percent of all psychotics, seventy-seven percent of the neurotics, eighty-four percent of those diagnosed situational maladjustment and forty-six percent of all personality disorders were classified as effective. Differential return-to-duty rates for the various diagnostic categories, combined with rather even success rates across diagnosis for those men returned to duty and the drop in the predictive power of diagnosis from a primary predictor to a relatively minor one, argue against diagnosis following disposition. In a study of naval outpatient settings by Fichman, et al (1973) diagnosis was again seen to be a powerful predictor of disposition ($R = .50$).

The demonstrated relationship between diagnosis and effectiveness and disposition indicates that in a clinical setting the psychiatrist is accurately reflecting the severity of the patient's psychological impairment. In this respect, the psychiatric diagnosis is functioning efficiently. With the knowledge that diagnosis performs a valuable service, it is important to investigate

how the individual clinician evaluates his patients and to know what variables he attends to in making a diagnostic decision.

Proposal - A Structural Analysis of Diagnosis

To study the diagnostic procedure in a meaningful way it is important to use an idiographic technique. George Kelly, in an effort to understand personal functioning, developed the Role Construct Repertory Test--Rep Test (Kelly, 1955). This test provides the user with the means to sample the important constructs an individual uses to give structure to his interpersonal environment. It is unusual because the person taking the test uses his cognitive structure and provides personal labels. Thus, in labels used and in factor structure generated it is uniquely idiographic.

The Rep Test sprang from Kelly's theories on personality. He proposed that men's actions are channeled by their anticipation of events. Men develop hypotheses about the world based on past experience. These hypotheses (constructions) are used both to understand the past and anticipate the future. Men differ in their construction of events, and may interpret the same events differently. In the interpretation of stimuli men develop a system of constructs involving ordinal relationships between constructs, constructs may overlap. To the extent that men use similar constructs to interpret experience, their psychological processes are similar (Kelly, 1955). This is not a complete exposition of Kelly's theory, but it provides a basis for understanding the process and products of the Rep Test.

In evaluating the usefulness of the Rep Test, it is important to know if the constructs generated are widely applied by the subject in his daily functioning. Field and Landfield (1961) examined the test-retest reliability of the Rep Test under three different conditions: (1) using the same figures for comparison, but requesting that the subject provide new constructs, (2) using

different figures and asking for constructs, (3) retaking the original test. Correlations generated by comparing agreements of the constructs produced were $r = .79$ for condition 1, $r = .79$ for condition 2 and $r = .80$ for condition 3. This experiment supports the hypothesis that the constructs produced by the Rep Test are stable entities.

Banister and Mair (1968) have written a comprehensive review of a number of measurement procedures classified as grid methods. The Rep Test is an example of this method. The subject is asked to compare and contrast selected figures. All figures are rated on all comparisons. These ratings are recorded in an n by n matrix. Generally, the grid method is the structural instrument which allows the individual examinee to develop his own content and psychological dimensions. It is the grid method which provides a fresh orientation to understanding the diagnostic decision of the clinician.

Edwards and Bennion (1974) used a grid method, in this case the Rep test, to assess clinical judgment. Eight psychologists, engaged in psychodiagnostic work, completed Rep Tests on their personal interactions. They reviewed a test protocol containing Rorschach, WAIS, MMPI, and Bender-Gestalt information. After reading the protocol, the psychologists were asked to rank order description statements about the patient. The descriptive statements were taken from the patient's and individual doctors' Rep Tests. Psychologists who chose statements from their own system first continued to see the patient in their own terms. Psychologists who chose a statement from the patient's system first were able to choose more data from the client's information.

In 1968, Allport proposed that the unreliability of individual prediction resided in the irrelevance of many of the dimensions applied to the individual. He advocated a morphogenic interest in the cognitive organization of one life

(Allport, 1968). The grid method represents a new approach to the problem of diagnosis consonant with Allport's demand for individuality. By applying a morphogenic technique to a personal process more reliable information about clinical judgment may be generated.

A Morphogenic Approach to Diagnosis

Method

Procedure. The grid method (Bannister & Mair, 1968) was used to collect data on 13 Navy psychiatrists at the San Diego Regional Medical Center and the Long Beach Regional Medical Center. The doctors had a mean time in practice past residency of 2.0 years with a range of 1 to 12 years.

The instrument consisted of two parts: (1) a role list overlay, and (2) the conceptual grid. See Figure 1. The overlay listed psychiatric diagnostic categories in a specified order. The grid was a 16 by 16 matrix with three circled comparisons in each row. Patient's names were supplied to the categories on the overlay. The overlay was used to provide material for conceptual sorts for each row of the grid.

Grids were administered in small groups of no more than 5 psychiatrists. Doctors were asked to list their patients on a 16 line overlay in the following order: psychosis, neurosis, personality disorder, situational maladjustment and alcoholism. This order was repeated for 16 lines. The role lists were adjusted to accommodate differences in individual caseloads where a doctor was not able to fill all classes. Adjustments were made by having psychiatrists substitute patients from other diagnostic categories when they had exhausted their cases from the requested categories and record the revised diagnosis on the proper overlay line. Figure 1 shows a partial grid and overlay.

Insert Figure 1 About Here (see pg 17a)

Data Analysis. Each doctor's grid was analyzed individually with the following procedure. Every patient's profile was compared with every other patient's profile. The number of matches of checks and voids was totaled. The number of matches was a measure of similarity between patients. The higher the measure, the greater the perceived similarity.

Within diagnostic class similarity was defined as the similarity measures for matched comparisons, e.g., psychotic - psychotic, neurotic - neurotic, personality disorder - personality disorder, situational maladjustment - situational maladjustment, and alcoholic - alcoholic. Between diagnostic class similarity was defined as the similarity measures for all comparisons between unmatched diagnostic categories, e.g., psychotic - neurotic, alcoholic - personality disorder.

Results

Between and within diagnostic class similarity measures were totaled and means computed across all doctors. The mean similarity measure within diagnostic classes ($X = 10.05$) was greater than the mean similarity measure between diagnostic classes ($X = 8.75$) ($t = 7.84$; $p .001$).

Mean similarity measures for between and within diagnostic classes and t values for the difference between groups for each doctor are presented in Table 1. For 8 out of 13 doctors, within diagnostic class similarity was significantly greater than between diagnostic class similarity.

INSERT TABLE 1 ABOUT HERE (see pg 18a)

Mean similarity scores for all mismatched diagnostic comparisons are presented in Table 2. In analyzing the comparisons, two groups emerged. Group 1 and group 2 had within group similarity and a significant difference between groups. In group 1, the neurotic - situational maladjustment comparison had the highest similarity score, and was significantly greater than all other mismatch

Table 1

Doctors' Mean Between and Within Diagnostic Class Similarity Score

Doctor#	Within Diagnostic Categories		Between Diagnostic Categories		+ Values
	\bar{X}	s.d.	\bar{X}	s.d.	
1	12.11	2.78	8.17	2.42	6.20
2	10.84	2.19	8.31	1.94	5.09
3	10.05	2.24	8.31	2.37	2.94
4	10.52	1.72	8.98	2.44	2.90
5	9.95	2.69	8.48	2.28	2.67
6	10.44	1.94	8.99	2.43	2.39
7	10.77	2.42	9.65	2.05	2.18
8	9.09	2.04	7.80	2.77	2.02
9	9.76	2.21	8.62	2.45	1.95
10	10.47	2.80	9.45	2.61	1.60
11	9.27	2.84	8.27	2.79	1.14
12	10.48	2.91	10.22	1.66	.58
13	8.00	3.29	8.50	2.85	-.81

comparisons except personality disorder - alcoholic and psychotic - alcoholic. In the three highest comparisons the alcoholic was seen as being similar to the psychotic and the personality disorder.

INSERT TABLE 2 ABOUT HERE (see pg 19a)

Comment

Use of a morphogenic technique revealed that a within-diagnostic class similarity was stable across all doctors. Psychiatrists used a diagnosis to label patients who were similar. Analysis of each individual profile indicated that some doctors were more successful than others in using diagnosis to indicate patient similarities. Some diagnostic classes appear to cause general confusion: Alcoholism with psychosis or personality disorder, and neurosis with situational maladjustment.

When doctors compare their patients on dimensions taken from the doctor's constructs, diagnosis becomes a useful tool to label patient similarities. The technique used in this study captures the properties of diagnosis that have made it a reliable predictor while exposing the individual variation in diagnostic functioning in real clinical settings.

Table 2
Diagnostic Similarity Scores for Mismatched Diagnoses

<u>Comparison</u>	Diagnostic Similarity Score	
	X	s.d.
Group 1		
Neurotic/Situational Maladjustment	9.69	2.18
Personality Disorder/Alcoholic	9.51	2.06
Psychotic/Alcoholic	9.27	2.43
Group 2		
Personality Disorder/Situational Maladjustment	8.73	2.26
Personality Disorder/Psychotic	8.71	2.71
Personality Disorder/Neurotic	8.70	2.45
Psychotic/Situational Maladjustment	8.46	2.38
Psychotic/Neurotic	8.30	2.43
Alcoholic/Situational Maladjustment	8.30	2.71
Alcoholic/Neurotic	7.95	2.87

References

- Allport, Gordon W. The Person in Psychology: Selected Essays. Beacon Press, Boston, 1968, 80-102.
- Arthur, R. J. and Gunderson, E. K. E. The prediction of diagnosis and disposition in naval hospitals. Journal of Clinical Psychology, 1966, 22, 459-464.
- Banister, D., and Mair J. M. M. The Evaluation of Personal Constructs. Academic Press, New York, 1968.
- Beenan, Folkert, Van Frankenhuisen, Jan H. and Veldkamp, Jacob G. The construction of a descriptive diagnostic system in psychiatry: First experience with a computer simulation. Behavioral Science, 1972, 17, 278-287.
- Edwards, Darrel and Bennion, Robert C. Intrapersonal constructs in understanding interpersonal behavior. Psychology, In press.
- Edwards, Darrel, Fichman, Susan, Bucky, Steven, and Berry, Newell H. A consideration of the state of the art of diagnosis in real clinical settings. Journal of Psychology, 1974, 87, 305-310.
- Edwards, Darrel, Gunderson, E. K. E., Brown, D. C., and Taylor, W. R. Clinical decisions and outcomes in a Navy psychiatric service. Military Medicine, 1973, 138, 430-432.
- Ellis, N. C. and Sells, S. B. An analysis of psychiatric diagnosis in a military mental hygiene clinic. Journal of Clinical Psychology, 1964, 20, 354-356.
- Fichman, Susan, Edwards, Darrel, and Berry, Newell H. Investigation of an industrial psychiatric decision: Disposition, diagnosis, attitudes, and motivation. Journal of Occupational Medicine, 1973, 15, 891-883.

- Fiedt, F. Harold. Comparison of visual content and auditory cues in interviewing. Journal of Consulting Psychology, 1955, 19, 407-416.
- Fjeld, S. P., Landfield, A. W. Personal construct consistency. Psychological Reports, 1961, 8, 127-129.
- Foulds, G. A. The reliability of psychiatric, and the validity of psychological diagnosis. Journal of Mental Science, 1955, 101, 851-862.
- Frank, George H. Psychiatric diagnosis: A review of research. Journal of General Psychology, 1969, 81, 157-176.
- Garfield, Sol L. Research on client variables in psychotherapy, from Bergin, Allen E., and Garfield, Sol L. Handbook of Psychotherapy and Behavior Change: An Empirical Analysis. John Wiley and Sons, Inc. New York, 1971, pp. 271-94.
- Gauron, Eugene F. and Dickinson, John K. Diagnostic decision making in psychiatry. Archives of General Psychiatry, 1966, 14, 225-232.
- Goldberg, Lewis R. Simple models or simple processes: Some research on clinical judgements. American Psychologist, 1968, 23(7), 483-496.
- Goldberg, Lewis R. and Werts, Charles E. The reliability of clinicians' judgements: A multitrait-multimethod approach. Journal of Consulting Psychology, 1966, 30, 199-206.
- Golden, Mark. Some effects of combining psychological tests on clinical inferences. Journal of Consulting Psychology, 1964, 28, 440-446.
- Grigg, Austin E. Experience of clinicians and speech characteristics and statements of clients are variables in clinical judgements. Journal of Consulting Psychology, 1958, 22, 315-319.
- Gunderson, E. K. Eric. The reliability of personality rating under varied assessment conditions. Journal of Clinical Psychology, 1965^a, 21, 161-164.
- Gunderson, E. K. Eric. Determinants of reliability in personality ratings. Journal of Clinical Psychology, 1965^b, 21, 164-169.

Harrison, Saul I., McDermott, John F., Schrager, Jules, and Showerman, Earl R.

Social status and child psychiatric practice: The influence of the clinicians' socioeconomic origin. American Journal of Psychiatry, 1970, 127, 652-658.

Hoffman, Paul L. The paramorphic representation of clinical judgment.

Psychological Bulletin, 1960, 57, 116-131.

Huff, Frederick W. and Friedman, Howard A. A study of the effects on agreement among clinicians of redundant and new information, confidence, and time available for assessments. Journal of General Psychology, 1967, 76, 49-57.

Hunt, William A., Jones, Nelson F., and Hunt, Edna B. Reliability of clinical judgment as a function of clinical experience. Journal of Clinical Psychology, 1957, 13, 377-378.

Johnston, Roy and McNeal, Benjamin F. Statistical versus clinical prediction: Length of neuropsychiatric hospital stay. Journal of Abnormal Psychology, 1967, 72, 335-340.

Kanfer, Frederick H. and Saslow, George, Behavioral diagnosis from Franks, Cyril M. Behavior Therapy: Appraisal and Status. McGraw-Hill, New York, 1969, 417-444.

Katz, Martin M. Cole, Johnathan O. and Lowery, Henri I. Studies of the diagnostic process: The influences of symptom perception, past experience, and ethnic background on diagnostic decision. American Journal of Psychiatry, 1969, 125, 937-946.

Kelly, George A. The Psychology of Personal Constructs Volume One: A Theory of Personality. W.W. Norton and Company, Inc., New York, 1955.

Kostian, Albert. A method for the empirical study of psychodiagnosis. Journal of Consulting Psychology, 1954, 18, 83-88.

- Lowinger, Paul and Dobie, Shirley. The attitudes of the psychiatrist about his patient. Comprehensive Psychiatry, 1968, 9, 627-632.
- Luft, Joseph. Implicit hypotheses and clinical prediction. Journal of Abnormal and Social Psychology, 1950, 45, 756-759.
- Mahrer, Alvin R., Thorp, Thomas, and Sternlicht, Irwin. The role of cues in psychodiagnosis. Journal of General Psychology, 1960, 62, 247-256.
- Meltzoff, Julian and Kornreich, Melvin. Research in Psychotherapy, Atherton Press, Inc. New York, 1970, pp. 204-336.
- Nathan, Peter E. A systems analytic model of diagnosis: V the diagnostic validity of disordered consciousness. Journal of Clinical Psychology, 1969, 25, 243-246.
- Nathan, Peter E., Robertson, P., and Adnberg, M. M. A systems analytic model of diagnosis: IV the diagnostic validity of abnormal affective behavior. Journal of Clinical Psychology, 1969, 25, 235-242.
- Oskamp, Stuart. The relationship of clinical experience and training methods to several criteria of clinical prediction. Psychological Monographs, 1962, 76, 28, Whole No. 547.
- Oskamp, Stuart. Overconfidence in case-study judgements. Journal of Consulting Psychology. 1965, 29, 261-265.
- Phelan, Joseph G. Rationale employed by clinical psychologists in diagnostic judgement. Journal of Clinical Psychology, 1964, 20, 454-458.
- Phelan, Joseph G. Use of matching method in measuring reliability of individual clinicians diagnostic judgement. Psychological Reports, 1965, 16, 491-496.
- Rapaport, David, Gill, Merton M., Schaefer, Roy. Diagnostic Psychological Testing. International Universities Press, Inc. New York, 1968.

- Schmidt, Hermann O. and Fonda, Charles P. The reliability of psychiatric diagnosis: A new look. Journal of Abnormal and Social Psychology, 1956, 52, 262-267.
- Shader, Richard I., Binstock, William A., Ohly, Jane I., and Scott, Dorothy. Biasing factors in diagnosis and disposition. Comprehensive Psychiatry, 1969, 10, 81-89.
- Stone, LeRoy A. Psychiatrists' prognostic judgements regarding functional psychotic disorders: A prognostic scale. Behavioral Science, 1966, 11, 115-120.
- Stone, LeRoy A. Judged prognosis: An attempt to apply a psychophysical approach to scale psychiatric opinion. Journal of Clinical Psychology, 1968, 24, 312-317.
- Stone, LeRoy A. Bases for psychiatric prognostic favorability judgement, psychophysical power functions. Behavioral Science, 1969, 14, 133-137.
- Stone, LeRoy A. A law of clinical judgement: A psychological mechanism based on the logic of psychophysics. Journal of Clinical Psychology, 1970, 26, 312-317.
- Zigler, Edward and Phillips, Leslie. Psychiatric diagnosis and symptomatology. Journal of Abnormal and Social Psychology, 1961, 63, 69-75.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 75-10	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) A Case For Psychiatric Diagnosis		5. TYPE OF REPORT & PERIOD COVERED Final report
7. AUTHOR(s) J. Susan Fichman and Darrel Edwards		6. PERFORMING ORG. REPORT NUMBER
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Health Research Center P.O. Box 85122 San Diego, CA 92138		8. CONTRACT OR GRANT NUMBER(s)
11. CONTROLLING OFFICE NAME AND ADDRESS Naval Medical Research and Development Command Bethesda, Maryland 20014		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS M4305.07-3013DXB5
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) Bureau of Medicine and Surgery Department of Navy Washington, D.C. 20375		12. REPORT DATE December 1974
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited. NAV HLT HRS CHC - 75-10		13. NUMBER OF PAGES 28
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report) Approved for public release; distribution unlimited.		15. SECURITY CLASS. (of this report) Unclassified
18. SUPPLEMENTARY NOTES		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) psychiatric diagnosis, patient evaluation, grid method, consistency of diagnostic class		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) Psychiatric diagnosis has been severely criticized by many clinicians and researchers. Studies in clinical judgment and diagnosis leading to this criticism were reviewed. Research designs were evaluated in terms of their relevance to diagnosis. Most projects were extremely narrow in scope and it was inappropriate to apply their conclusions to the process of diagnosis. Research evaluating diagnosis in real clinical settings showed that diagnosis accurately and consistently assessed the patients' capacity for successful		

DD FORM 1473
1 JAN 73EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

391 642

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

post-hospitalization readjustment. In an effort to evaluate diagnosis as practiced in the clinical setting, the grid method was used. Psychiatrists rated patients from their practice, generating their own dimensions for comparison. Patients within a diagnostic category were consistently evaluated as being more similar to each other than patients from different diagnostic categories. The use of this highly individual technique assured that even analysis of grouped data would not obscure the basic strength of the diagnostic process.

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)